The Joy of X

-Steven Strogatz

2 Rock Groups

Like anything else, arithmetic has its serious side and its playful side.

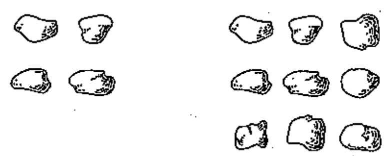
The serious side is what we all learned in school: how to work with columns of numbers, adding them, subtracting them, grinding them through the spreadsheet calculations needed for tax returns and year-end reports. This side of arithmetic is important, practical, and—for many people—joyless.

The playful side of arithmetic is a lot less familiar, unless you were trained in the ways of advanced mathematics. Yet there's nothing inherently advanced about it. It's as natural as a child's curiosity.

In his book *A Mathematician's Lament*, Paul Lockhart advocates an educational approach in which numbers are treated more concretely than usual: he asks us to imagine them as groups of rocks. For example, 6 corresponds to a group of rocks like this:

You probably don't see anything striking here, and that's right —unless we make further demands on numbers, they all look pretty much the same. Our chance to be creative comes in what we ask of them.

For instance, let's focus on groups having between 1 and 10 rocks in them, and ask which of these groups can be rearranged into square patterns. Only two of them can: the group with 4 and the group with 9. And that's because $4 = 2 \times 2$ and $9 = 3 \times 3$; we get these numbers by squaring some other number (actually making a square shape).



A less stringent challenge is to identify groups of rocks that can be neatly organized into a rectangle with exactly two rows that come out even. That's possible as long as there are 2, 4, 6, 8, or 10 rocks; the number has to be even. If we try to coerce any of the other numbers from 1 to 10—the odd numbers—into two rows, they always leave an odd bit sticking out.

Still, all is not lost for these misfit numbers. If we add two of them together, their protuberances match up and their sum comes out even; Odd + Odd = Even.

If we loosen the rules still further to admit numbers greater than 10 and allow a rectangular pattern to have more than two rows of rocks, some odd numbers display a talent for making these larger rectangles. For example, the number 15 can form a 3×5 rectangle:

So 15, although undeniably odd, at least has the consolation of being a composite number—it's composed of three rows of five rocks each. Similarly, every other entry in the multiplication table yields its own rectangular rock group.

Yet some numbers, like 2, 3, 5, and 7, truly are hopeless. None of them can form any sort of rectangle at all, other than a simple line of rocks (a single row). These strangely inflexible numbers are the famous prime numbers.

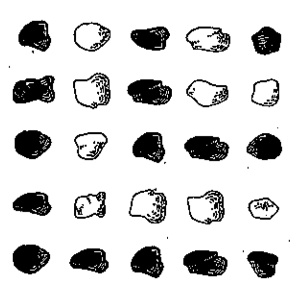
So we see that numbers have quirks of structure that endow them with personalities. But to see the full range of their behavior, we need to go beyond individual numbers and watch what happens when they interact.

For example, instead of adding just two odd numbers together, suppose we add all the consecutive odd numbers, starting from 1:

$$1 + 3 = 4$$
$$1 + 3 + 5 = 9$$
$$1 + 3 + 5 + 7 = 16$$
$$1 + 3 + 5 + 7 + 9 = 25.$$

The sums above, remarkably, always turn out to be perfect squares. (We saw 4 and 9 in the square patterns discussed earlier, and $16 = 4 \times 4$, and $25 = 5 \times 5$.) A quick check shows that this rule keeps working for larger and larger odd numbers; it apparently holds all the way out to infinity. But what possible connection could there be between odd numbers, with their ungainly appendages, and the classically symmetrical numbers that form squares? By arranging our rocks in the right way, we can make this surprising link seem obvious—the hallmark of an elegant proof.

The key is to recognize that odd numbers can make L-shapes, with their protuberances cast off into the corner. And when you stack successive L-shapes together, you get a square!



This style of thinking appears in another recent book, though for altogether different literary reasons. In Yoko Ogawa's charming novel *The Housekeeper and the Professor*, an astute but uneducated young woman with a ten-year-old son is hired to take care of an elderly mathematician who has suffered a traumatic brain injury that leaves him with only eighty minutes of short-term memory. Adrift in the present, and alone in his shabby cottage with nothing but his numbers, the Professor tries to connect with the Housekeeper the only way he knows how: by inquiring about her shoe size or birthday and making mathematical small talk about her statistics. The Professor also takes a special liking to the Housekeeper's son, whom he calls Root, because the flat top of the boy's head reminds him of the square root symbol, $\sqrt{\ }$.

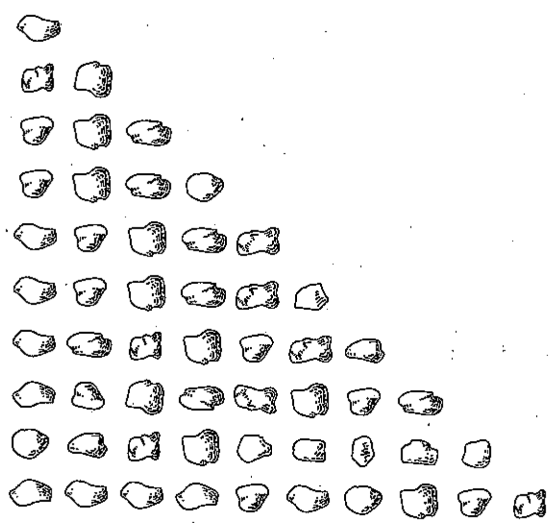
One day the Professor gives Root a little puzzle: Can he find the sum of all the numbers from 1 to 10? After Root carefully adds the numbers and returns with the answer (55), the Professor asks him to find a better way. Can he find the answer *without* adding the numbers? Root kicks the chair and shouts, "That's not fair!"

But little by little the Housekeeper gets drawn into the world of numbers, and she secretly starts exploring the puzzle herself. "I'm not sure why I became so absorbed in a child's math problem with no practical value," she says. "At first, I was conscious of wanting to please the Professor, but gradually that feeling faded and I realized it had become a battle between the problem and me. When I woke in the morning, the equation was waiting:

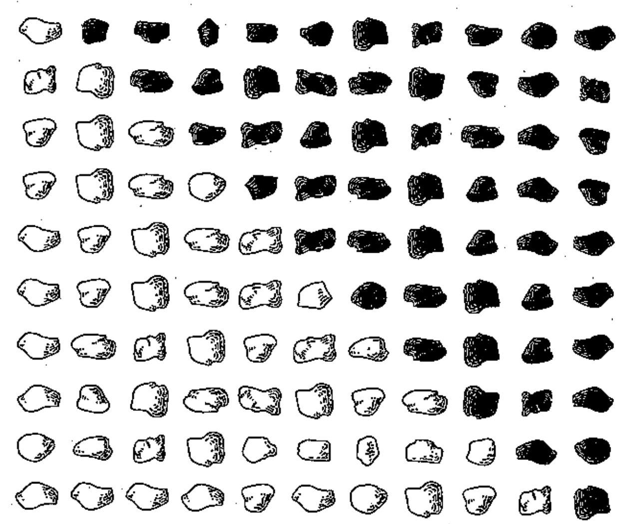
$$1 + 2 + 3 + \ldots + 9 + 10 = 55$$

and it followed me all through the day, as though it had burned itself into my retina and could not be ignored."

There are several ways to solve the Professor's problem (see how many you can find). The Professor himself gives an argument along the lines we developed above. He interprets the sum from 1 to 10 as a triangle of rocks, with 1 rock in the first row, 2 in the second, and so on, up to 10 rocks in the tenth row:



By its very appearance this picture gives a clear sense of negative space. It seems only half complete. And that suggests a creative leap. If you copy the triangle, flip it upside down, and add it as the missing half to what's already there, you get something much simpler: a rectangle with ten rows of 11 rocks each, for a total of 110.



Since the original triangle is half of this rectangle, the desired sum must be half of 110, or 55.

Looking at numbers as groups of rocks may seem unusual, but actually it's as old as math itself. The word "calculate" reflects that legacy—it comes from the Latin word *calculus*, meaning a pebble used for counting. To enjoy working with numbers you don't have to be Einstein (German for "one stone"), but it might help to have rocks in your head.

3 The Enemy of My Enemy

IT'S TRADITIONAL TO teach kids subtraction right after addition. That makes sense—the same facts about numbers get used in both, though in reverse. And the black art of borrowing, so crucial to successful subtraction, is only a little more baroque than that of carrying, its counterpart for addition. If you can cope with calculating 23 + 9, you'll be ready for 23 – 9 soon enough.

At a deeper level, however, subtraction raises a much more disturbing issue, one that never arises with addition. Subtraction can generate negative numbers. If I try to take 6 cookies away from you but you have only 2, I can't do it—except in my mind, where you now have negative 4 cookies, whatever that means.

Subtraction forces us to expand our conception of what numbers are. Negative numbers are a lot more abstract than positive numbers—you can't see negative 4 cookies and you certainly can't eat them—but you can think about them, and you *have* to, in all aspects of daily life, from debts and overdrafts to contending with freezing temperatures and parking garages.

Still, many of us haven't quite made peace with negative

numbers. As my colleague Andy Ruina has pointed out, people have concocted all sorts of funny little mental strategies to sidestep the dreaded negative sign. On mutual fund statements, losses (negative numbers) are printed in red or nestled in parentheses with nary a negative sign to be found. The history books tell us that Julius Caesar was born in 100 B.C., not –100. The subterranean levels in a parking garage often have designations like B1 and B2. Temperatures are one of the few exceptions: folks do say, especially here in Ithaca, New York, that it's –5 degrees outside, though even then, many prefer to say 5 below zero. There's something about that negative sign that just looks so unpleasant, so . . . negative.

Perhaps the most unsettling thing is that a negative times a negative is a positive. So let me try to explain the thinking behind that.

How should we define the value of an expression like -1×3, where we're multiplying a negative number by a positive number? Well, just as 1×3 means $1 + 1 + 1$, the natural definition for -1×3 is $(-1) + (-1) + (-1)$, which equals -3. This should be obvious in terms of money: if you owe me $1 a week, after three weeks you're $3 in the hole.

From there it's a short hop to see why a negative times a negative should be a positive. Take a look at the following string of equations:

$$\begin{aligned}
-1 \times 3 &= -3 \\
-1 \times 2 &= -2 \\
-1 \times 1 &= -1 \\
-1 \times 0 &= 0 \\
-1 \times -1 &= ?
\end{aligned}$$

Now look at the numbers on the far right and notice their

orderly progression: –3, –2, –1, 0, . . . At each step, we're adding 1 to the number before it. So wouldn't you agree the next number should logically be 1?

That's one argument for why $(-1) \times (-1) = 1$. The appeal of this definition is that it preserves the rules of ordinary arithmetic; what works for positive numbers also works for negative numbers.

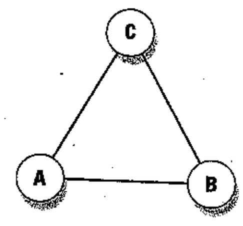
But if you're a hard-boiled pragmatist, you may be wondering if these abstractions have any parallels in the real world. Admittedly, life sometimes seems to play by different rules. In conventional morality, two wrongs don't make a right. Likewise, double negatives don't always amount to positives; they can make negatives more intense, as in "I can't get no satisfaction." (Actually, languages can be very tricky in this respect. The eminent linguistic philosopher J. L. Austin of Oxford once gave a lecture in which he asserted that there are many languages in which a double negative makes a positive but none in which a double positive makes a negative—to which the Columbia philosopher Sidney Morgenbesser, sitting in the audience, sarcastically replied, "Yeah, yeah.")

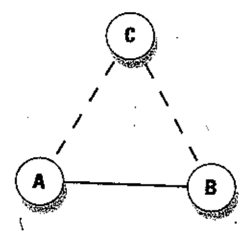
Still, there are plenty of cases where the real world does mirror the rules of negative numbers. One nerve cell's firing can be inhibited by the firing of a second nerve cell. If that second nerve cell is then inhibited by a third, the first cell can fire again. The indirect action of the third cell on the first is tantamount to excitation; a chain of two negatives makes a positive. Similar effects occur in gene regulation: a protein can turn a gene on by blocking another molecule that was repressing that stretch of DNA.

Perhaps the most familiar parallel occurs in the social and political realms as summed up by the adage "The enemy of my enemy is my friend." This truism, and related ones about the

friend of my enemy, the enemy of my friend, and so on, can be depicted in relationship triangles.

The corners signify people, companies, or countries, and the sides connecting them signify their relationships, which can be positive (friendly, shown here as solid lines) or negative (hostile, shown as dashed lines).

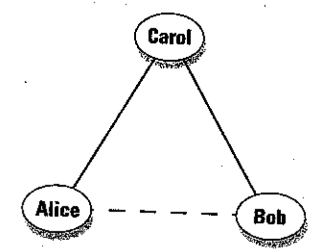




Social scientists refer to triangles like the one on the left, with all sides positive, as balanced—there's no reason for anyone to change how he feels, since it's reasonable to like your friends' friends. Similarly, the triangle on the right, with two negatives and a positive, is considered balanced because it causes no dissonance; even though it allows for hostility, nothing cements a friendship like hating the same person.

Of course, triangles can also be unbalanced. When three mutual enemies size up the situation, two of them—often the two with the least animosity toward each other—may be tempted to join forces and gang up on the third.

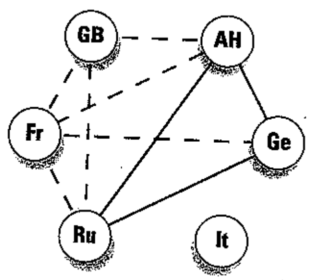
Even more unbalanced is a triangle with a single negative relationship. For instance, suppose Carol is friendly with both Alice and Bob, but Bob and Alice despise each other. Perhaps they were once a couple but suffered a nasty breakup, and each is now badmouthing the other to ever-loyal Carol. This causes psychological stress all around. To restore balance, either Alice and Bob have to reconcile or Carol has to choose a side.



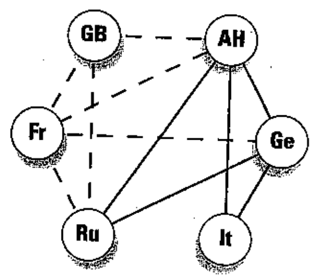
In all these cases, the logic of balance matches the logic of multiplication. In a balanced triangle, the sign of the product of any two sides, positive or negative, always agrees with the sign of the third. In unbalanced triangles, this pattern is broken.

Leaving aside the verisimilitude of the model, there are interesting questions here of a purely mathematical flavor. For example, in a close-knit network where everyone knows everyone, what's the most stable state? One possibility is a nirvana of goodwill, where all relationships are positive and all triangles within the network are balanced. But surprisingly, there are other states that are equally stable. These are states of intractable conflict, with the network split into two hostile factions (of arbitrary sizes and compositions). All members of one faction are friendly with one another but antagonistic toward everybody in the other faction. (Sound familiar?) Perhaps even more surprisingly, these polarized states are the *only* states as stable as nirvana. In particular, no three-party split can have all its triangles balanced.

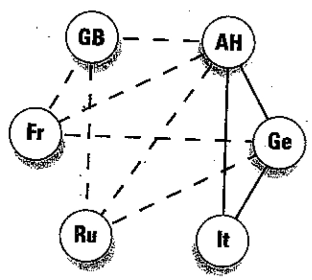
Scholars have used these ideas to analyze the run-up to World War I. The diagram that follows shows the shifting alliances among Great Britain, France, Russia, Italy, Germany, and Austria-Hungary between 1872 and 1907.



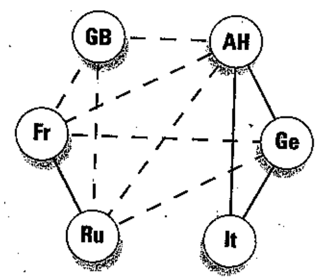
Three Emperors' League
1872–81



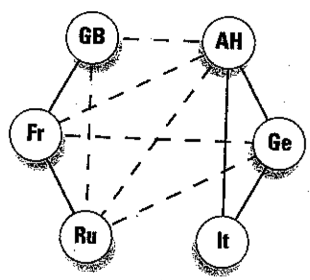
Triple Alliance 1882



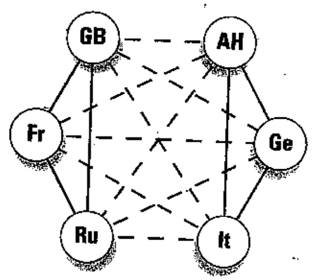
German-Russian Lapse 1890



French-Russian Alliance
1891–94



Entente Cordiale 1904



British-Russian Alliance 1907

The first five configurations were all unbalanced, in the sense that they each contained at least one unbalanced triangle. The resultant dissonance tended to push these nations to realign themselves, triggering reverberations elsewhere in the network. In the final stage, Europe had split into two implacably opposed blocs—technically balanced, but on the brink of war.

The point is not that this theory is powerfully predictive. It isn't. It's too simple to account for all the subtleties of geopolitical dynamics. The point is that some part of what we observe is due to nothing more than the primitive logic of "the enemy of my enemy," and *this* part is captured perfectly by the multiplication of negative numbers. By sorting the meaningful from the generic, the arithmetic of negative numbers can help us see where the real puzzles lie.

4 Commuting

EVERY DECADE OR so a new approach to teaching math comes along and creates fresh opportunities for parents to feel inadequate. Back in the 1960s, my parents were flabbergasted by their inability to help me with my second-grade homework. They'd never heard of base 3 or Venn diagrams.

Now the tables have turned. "Dad, can you show me how to do these multiplication problems?" *Sure*, I thought, until the headshaking began. "No, Dad, that's not how we're supposed to do it. That's the old-school method. Don't you know the lattice method? No? Well, what about partial products?"

These humbling sessions have prompted me to revisit multiplication from scratch. And it's actually quite subtle, once you start to think about it.

Take the terminology. Does "seven times three" mean "seven added to itself three times"? Or "three added to itself seven times"?

In some cultures the language is less ambiguous. A friend of mine from Belize used to recite his times tables like this: "Seven ones are seven, seven twos are fourteen, seven threes are twenty-one," and so on. This phrasing makes it clear that the first number is the multiplier; the second number is the thing being multiplied. It's the same convention as in Lionel Richie's

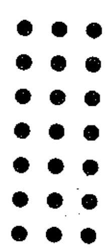
immortal lyrics "She's once, twice, three times a lady." ("She's a lady times three" would never have been a hit.)

Maybe all this semantic fuss strikes you as silly, since the order in which numbers are multiplied doesn't matter anyway: $7 \times 3 = 3 \times 7$. Fair enough, but that begs the question I'd like to explore in some depth here: Is this commutative law of multiplication, $a \times b = b \times a$, really so obvious? I remember being surprised by it as a child; maybe you were too.

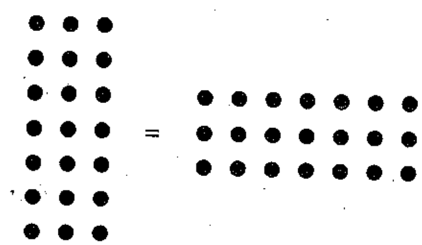
To recapture the magic, imagine not knowing what 7×3 equals. So you try counting by sevens: 7, 14, 21. Now turn it around and count by threes instead: 3, 6, 9, . . . Do you feel the suspense building? So far, none of the numbers match those in the sevens list, but keep going . . . 12, 15, 18, and then, bingo, 21!

My point is that if you regard multiplication as being synonymous with repeated counting by a certain number (or, in other words, with repeated addition), the commutative law isn't transparent.

But it becomes more intuitive if you conceive of multiplication *visually*. Think of 7×3 as the number of dots in a rectangular array with seven rows and three columns.



If you turn the array on its side, it transforms into three rows and seven columns—and since rotating the picture doesn't change the number of dots, it must be true that $7 \times 3 = 3 \times 7$.



$7 \times 3 \quad = \quad 3 \times 7$

Yet strangely enough, in many real-world situations, especially where money is concerned, people seem to forget the commutative law, or don't realize it applies. Let me give you two examples.

Suppose you're shopping for a new pair of jeans. They're on sale for 20 percent off the sticker price of $50, which sounds like a bargain, but keep in mind that you also have to pay the 8 percent sales tax. After the clerk finishes complimenting you on the flattering fit, she starts ringing up the purchase but then pauses and whispers, in a conspiratorial tone, "Hey, let me save you some money. I'll apply the tax first, and then take twenty percent off the total, so you'll get more money back. Okay?"

But something about that sounds fishy to you. "No thanks," you say. "Could you please take the twenty percent off first, then apply the tax to the sale price? That way, I'll pay less tax."

Which way is a better deal for you? (Assume both are legal.)

When confronted with a question like this, many people approach it *additively*. They work out the tax and the discount under both scenarios, and then do whatever additions or subtractions are necessary to find the final price. Doing things the clerk's way, you reason, would cost you $4 in tax (8 percent of the sticker price of $50). That would bring your total to $54.

Then applying the 20 percent discount to \$54 gives you \$10.80 back, so you'd end up paying \$54 minus \$10.80, which equals \$43.20. Whereas under your scenario, the 20 percent discount would be applied first, saving you \$10 off the \$50 sticker price. Then the 8 percent tax on that reduced price of \$40 would be \$3.20, so you'd still end up paying \$43.20. Amazing!

But it's merely the commutative law in action. To see why, think *multiplicatively*, not additively. Applying an 8 percent tax followed by a 20 percent discount amounts to multiplying the sticker price by 1.08 and then multiplying that result by 0.80. Switching the order of tax and discount reverses the multiplication, but since $1.08 \times 0.80 = 0.80 \times 1.08$, the final price comes out the same.

Considerations like these also arise in larger financial decisions. Is a Roth 401(k) better or worse than a traditional retirement plan? More generally, if you have a pile of money to invest and you have to pay taxes on it at some point, is it better to take the tax bite at the beginning of the investment period, or at the end?

Once again, the commutative law shows it's a wash, all other things being equal (which, sadly, they often aren't). If, for both scenarios, your money grows by the same factor and gets taxed at the same rate, it doesn't matter whether you pay the taxes up front or at the end.

Please don't mistake this mathematical remark for financial advice. Anyone facing these decisions in real life needs to be aware of many complications that muddy the waters: Do you expect to be in a higher or lower tax bracket when you retire? Will you max out your contribution limits? Do you think the government will change its policies about the tax-exempt status of withdrawals by the time you're ready to take the money out? Leaving all this aside (and don't get me wrong, it's all important; I'm just trying to focus here on a simpler mathematical issue), my basic point is that the commutative law is relevant to the analysis of such decisions.

You can find heated debates about this on personal finance sites on the Internet. Even after the relevance of the commutative law has been pointed out, some bloggers don't accept it. It's that counterintuitive.

Maybe we're wired to doubt the commutative law because in daily life, it usually matters what you do first. You can't have your cake and eat it too. And when taking off your shoes and socks, you've got to get the sequencing right.

The physicist Murray Gell-Mann came to a similar realization one day when he was worrying about his future. As an undergraduate at Yale, he desperately wanted to stay in the Ivy League for graduate school. Unfortunately Princeton rejected his application. Harvard said yes but seemed to be dragging its feet about providing the financial support he needed. His best option, though he found it depressing, was MIT. In Gell-Mann's eyes, MIT was a grubby technological institute, beneath his rarefied taste. Nevertheless, he accepted the offer. Years later he would explain that he had contemplated suicide at the time but decided against it once he realized that attending MIT and killing himself didn't commute. He could always go to MIT and commit suicide later if he had to, but not the other way around.

Gell-Mann had probably been sensitized to the importance of non-commutativity. As a quantum physicist he would have been acutely aware that at the deepest level, nature disobeys the commutative law. And it's a good thing, too. For the failure of commutativity is what makes the world the way it is. It's why matter is solid, and why atoms don't implode.

Specifically, early in the development of quantum mechan-

ics, Werner Heisenberg and Paul Dirac had discovered that nature follows a curious kind of logic in which $p \times q \neq q \times p$, where p and q represent the momentum and position of a quantum particle. Without that breakdown of the commutative law, there would be no Heisenberg uncertainty principle, atoms would collapse, and nothing would exist.

That's why you'd better mind your *p*'s and *q*'s. And tell your kids to do the same.